

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/988,202

DATE: 01/15/2002

TIME: 19:00:49

Input Set : N:\Crf3\RULE60\09988202.raw

Output Set: N:\CRF3\01152002\I988202.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: BARBEYRON, Tristan

6 POTIN, Philippe

7 RICHARD, Christophe

8 HENRISSAT, Bernard

9 YVIN, Jean-Claude

10 KLOAREG, Bernard

13 (ii) TITLE OF INVENTION: Glycolyse hydrolase genes and their

14 use for producing enzymes for the biodegradation of
15 carrageenans

17 (iii) NUMBER OF SEQUENCES: 8

19 (iv) CORRESPONDENCE ADDRESS:

20 (A) ADDRESSEE: DENNISON, MESEROLE, SCHEINER & SCHULTZ

21 (B) STREET: 612 Crystal Square 4, 1745 Jefferson Davis
22 Highway

23 (C) CITY: ARLINGTON

24 (D) STATE: VIRGINIA

25 (E) COUNTRY: U.S.A.

26 (F) ZIP: 22202

28 (v) COMPUTER READABLE FORM:

29 (A) MEDIUM TYPE: Floppy disk

30 (B) COMPUTER: IBM PC compatible

31 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

32 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

34 (vi) CURRENT APPLICATION DATA:

C--> 35 (A) APPLICATION NUMBER: US/09/988,202

C--> 36 (B) FILING DATE: 19-Nov-2001

37 (C) CLASSIFICATION:

43 (vii) PRIOR APPLICATION DATA:

40 (A) APPLICATION NUMBER: 09/269,731

41 (B) FILING DATE:

44 (A) APPLICATION NUMBER: FR 96 12204

45 (B) FILING DATE: 07-OCT-1996

47 (viii) ATTORNEY/AGENT INFORMATION:

48 (A) NAME: IRA SCHULTZ

49 (B) REGISTRATION NUMBER:

50 (C) REFERENCE/DOCKET NUMBER:

52 (ix) TELECOMMUNICATION INFORMATION:

53 (A) TELEPHONE: (703) 412-1155

54 (B) TELEFAX: (703) 412-1161

56 (2) INFORMATION FOR SEQ ID NO: 1:

58 (i) SEQUENCE CHARACTERISTICS:

59 (A) LENGTH: 2085 base pairs

60 (B) TYPE: nucleic acid

61 (C) STRANDEDNESS: single

62 (D) TOPOLOGY: linear

ENTERED

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/988,202

TIME: 19:00:49

Input Set : N:\Crif3\RULE60\09988202.raw

Output Set: N:\CRF3\01152002\I988202.raw

```

64      (ii) MOLECULE TYPE: DNA (genomic)
66      (iii) HYPOTHETICAL: NO
70      (ix) FEATURE:
71          (A) NAME/KEY: CDS
72          (B) LOCATION:join(211..1683, 1880..2083)
74      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
76 AAGCTTTCCG ATTCTATCAT CGAAGTCATA GGAGTGGGTA AACAAAAAAG CATGAAACTA      60
77 GCTTTTAAAT ATACAGACTT TCAATATAGG TCGCACACAA TATTAACGAA TAAATAAGCA      120
78 AATCATATAC ATAATCATTG CTTTAAATAT GTTTTAATAC AGATATAAAC ATAGTATGTT      180
79 TGTGTTTTTG GTATCTATCG GAGTGAAAAC ATG CGC TTA TAT TTT AGA AAG TTG      234
80                                     Met Arg Leu Tyr Phe Arg Lys Leu
81                                     1           5
83 TGG TTA ACA AAT TTA TTT TTA GGC GGA GCA CTG GCC TCT TCA GCT GCG      282
84 Trp Leu Thr Asn Leu Phe Leu Gly Gly Ala Leu Ala Ser Ser Ala Ala
85      10           15           20
87 ATA GGG GCT GTC TCC CCC AAG ACT TAT AAG GAC GCA GAT TTT TAT GTT      330
88 Ile Gly Ala Val Ser Pro Lys Thr Tyr Lys Asp Ala Asp Phe Tyr Val
89      25           30           35           40
91 GCC CCT ACT CAA CAA GAT GTT AAC TAT GAT TTA GTT GAT GAT TTT GGC      378
92 Ala Pro Thr Gln Gln Asp Val Asn Tyr Asp Leu Val Asp Asp Phe Gly
93      45           50           55
95 GCT AAT GGA AAC GAC ACT AGT GAT GAC AGT AAT GCT TTA CAA AGA GCA      426
96 Ala Asn Gly Asn Asp Thr Ser Asp Asp Ser Asn Ala Leu Gln Arg Ala
97      60           65           70
99 ATT AAT GCT ATT AGT AGA AAA CCG AAT GGG GGC ACT TTA CTA ATA CCG      474
100 Ile Asn Ala Ile Ser Arg Lys Pro Asn Gly Gly Thr Leu Leu Ile Pro
101      75           80           85
103 AAT GGA ACT TAC CAT TTC CTC GGC ATA CAG ATG AAG TCG AAC GTA CAC      522
104 Asn Gly Thr Tyr His Phe Leu Gly Ile Gln Met Lys Ser Asn Val His
105      90           95           100
107 ATC CGT GTT GAG AGT GAC GTG ATA ATC AAG CCA ACG TGG AAT GGG GAT      570
108 Ile Arg Val Glu Ser Asp Val Ile Ile Lys Pro Thr Trp Asn Gly Asp
109      105           110           115           120
111 GGC AAA AAC CAC CGA CTA TTT GAA GTT GGC GTA AAC AAT ATT GTA AGA      618
112 Gly Lys Asn His Arg Leu Phe Glu Val Gly Val Asn Asn Ile Val Arg
113      125           130           135
115 AAC TTC AGC TTT CAA GGG TTA GGA AAC GGT TTT TTG GTG GAT TTT AAA      666
116 Asn Phe Ser Phe Gln Gly Leu Gly Asn Gly Phe Leu Val Asp Phe Lys
117      140           145           150
119 GAT TCT CGC GAC AAA AAC TTA GCT GTT TTT AAG TTA GGC GAT GTT AGA      714
120 Asp Ser Arg Asp Lys Asn Leu Ala Val Phe Lys Leu Gly Asp Val Arg
121      155           160           165
123 AAT TAC AAA ATT TCC AAT TTT ACC ATT GAT GAT AAT AAA ACG ATA TTT      762
124 Asn Tyr Lys Ile Ser Asn Phe Thr Ile Asp Asp Asn Lys Thr Ile Phe
125      170           175           180
127 GCC TCA ATT TTA GTG GAC GTA ACA GAA CGT AAT GGG CGG TTA CAT TGG      810
128 Ala Ser Ile Leu Val Asp Val Thr Glu Arg Asn Gly Arg Leu His Trp
129      185           190           195           200
131 TCG CGT AAT GGA ATT ATC GAA AGA ATA AAA CAA AAT AAC GCT TTG TTC      858

```

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/988,202

TIME: 19:00:49

Input Set : N:\Crif3\RULE60\09988202.raw

Output Set: N:\CRF3\01152002\I988202.raw

132	Ser	Arg	Asn	Gly	Ile	Ile	Glu	Arg	Ile	Lys	Gln	Asn	Asn	Ala	Leu	Phe	
133					205					210					215		
135	GGC	TAC	GGC	CTT	ATT	CAA	ACC	TAT	GGC	GCA	GAT	AAT	ATT	TTG	TTT	AGG	906
136	Gly	Tyr	Gly	Leu	Ile	Gln	Thr	Tyr	Gly	Ala	Asp	Asn	Ile	Leu	Phe	Arg	
137				220						225					230		
139	AAC	CTC	CAT	TCG	GAA	GGC	GGA	ATT	GCG	TTA	CGG	ATG	GAA	ACT	GAC	AAC	954
140	Asn	Leu	His	Ser	Glu	Gly	Gly	Ile	Ala	Leu	Arg	Met	Glu	Thr	Asp	Asn	
141				235						240					245		
143	TTA	CTT	ATG	AAA	AAT	TAT	AAG	CAA	GGC	GGA	ATA	AGA	AAC	ATC	TTT	GCT	1002
144	Leu	Leu	Met	Lys	Asn	Tyr	Lys	Gln	Gly	Gly	Ile	Arg	Asn	Ile	Phe	Ala	
145				250						255					260		
147	GAT	AAT	ATC	AGA	TGT	AGC	AAA	GGA	CTT	GCG	GCG	GTC	ATG	TTT	GGC	CCA	1050
148	Asp	Asn	Ile	Arg	Cys	Ser	Lys	Gly	Leu	Ala	Ala	Val	Met	Phe	Gly	Pro	
149	265															280	
151	CAT	TTT	ATG	AAG	AAT	GGA	GAT	GTG	CAA	GTG	ACC	AAT	GTC	AGC	TCA	GTT	1098
152	His	Phe	Met	Lys	Asn	Gly	Asp	Val	Gln	Val	Thr	Asn	Val	Ser	Ser	Val	
153					285										295		
155	AGT	TGC	GGT	TCG	GCT	GTA	CGA	AGT	GAT	AGT	GGA	TTT	GTC	GAA	CTC	TTT	1146
156	Ser	Cys	Gly	Ser	Ala	Val	Arg	Ser	Asp	Ser	Gly	Phe	Val	Glu	Leu	Phe	
157				300						305					310		
159	AGC	CCG	ACA	GAC	GAA	GTA	CAT	ACG	CGT	CAA	AGT	TGG	AAA	CAA	GCC	GTT	1194
160	Ser	Pro	Thr	Asp	Glu	Val	His	Thr	Arg	Gln	Ser	Trp	Lys	Gln	Ala	Val	
161				315						320					325		
163	GAA	AGT	AAA	TTG	GGC	CGA	GGG	TGT	GCG	CAA	ACC	CCT	TAT	GCT	AGA	GGT	1242
164	Glu	Ser	Lys	Leu	Gly	Arg	Gly	Cys	Ala	Gln	Thr	Pro	Tyr	Ala	Arg	Gly	
165				330						335					340		
168	AAT	GGT	GGT	ACA	CGG	TGG	GCG	GCT	CGC	GTA	ACA	CAA	AAA	GAC	GCG	TGT	1290
169	Asn	Gly	Gly	Thr	Arg	Trp	Ala	Ala	Arg	Val	Thr	Gln	Lys	Asp	Ala	Cys	
170	345															360	
172	TTA	GAT	AAA	GCA	AAA	CTG	GAA	TAT	GGA	ATA	GAG	CCT	GGT	TCA	TTT	GGC	1338
173	Leu	Asp	Lys	Ala	Lys	Leu	Glu	Tyr	Gly	Ile	Glu	Pro	Gly	Ser	Phe	Gly	
174					365										375		
176	ACG	GTT	AAA	GTC	TTT	GAT	GTT	ACA	GCG	CGT	TTT	GGT	TAT	AAC	GCA	GAT	1386
177	Thr	Val	Lys	Val	Phe	Asp	Val	Thr	Ala	Arg	Phe	Gly	Tyr	Asn	Ala	Asp	
178					380					385					390		
180	CTT	AAA	CAG	GAC	CAG	CTA	GAC	TAC	TTT	TCT	ACA	TCC	AAC	CCT	ATG	TGC	1434
181	Leu	Lys	Gln	Asp	Gln	Leu	Asp	Tyr	Phe	Ser	Thr	Ser	Asn	Pro	Met	Cys	
182				395						400					405		
184	AAG	CGT	GTA	TGC	CTT	CCT	ACA	AAA	GAA	CAA	TGG	AGT	AAG	CAA	GGC	CAA	1482
185	Lys	Arg	Val	Cys	Leu	Pro	Thr	Lys	Glu	Gln	Trp	Ser	Lys	Gln	Gly	Gln	
186				410						415					420		
188	ATT	TAC	ATT	GGT	CCG	TCA	TTA	GCT	GCA	GTA	ATT	GAT	ACC	ACA	CCT	GAA	1530
189	Ile	Tyr	Ile	Gly	Pro	Ser	Leu	Ala	Ala	Val	Ile	Asp	Thr	Thr	Pro	Glu	
190	425															440	
192	ACT	TCA	AAA	TAC	GAT	TAT	GAT	GTG	AAA	ACT	TTT	AAC	GTC	AAA	AGA	ATA	1578
193	Thr	Ser	Lys	Tyr	Asp	Tyr	Asp	Val	Lys	Thr	Phe	Asn	Val	Lys	Arg	Ile	
194					445										455		
196	AAT	TTT	CCT	GTA	AAT	TCA	CAC	AAG	ACT	ATC	GAC	ACG	AAT	ACT	GAA	AGT	1626
197	Asn	Phe	Pro	Val	Asn	Ser	His	Lys	Thr	Ile	Asp	Thr	Asn	Thr	Glu	Ser	

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/09/988,202

TIME: 19:00:49

Input Set : N:\Crif3\RULE60\09988202.raw

Output Set: N:\CRF3\01152002\I988202.raw

```

198          460          465          470
200 AGC CGT GTC TGC AAT TAT TAC GGT ATG TCC GAA TGC TCC AGC AGT CGA      1674
201 Ser Arg Val Cys Asn Tyr Tyr Gly Met Ser Glu Cys Ser Ser Ser Arg
202          475          480          485
204 TGG GAG CGA TAGATTAAGC CGCTATATTC ATTTACTAGG TAAAACTTCA      1723
205 Trp Glu Arg
206          490
208 AGCCGCATTC GAAGAACTAT CGAACGCGGC TTTTGTGTTA AGAGCGCCTA TGA CTCAGTA      1783
209 TATTTTGTAT AAATATAATT TTACATCTTG TTAAAGTAAA CATCATATGT TTATATAGGT      1843
210 GCAATCTAAT TTGTTAATAT AGTGTTGGAG ATAGGT ATG AAA GGT GTT TCT ACG      1897
211          Met Lys Gly Val Ser Thr
212          495
215 AAA AAT GCT CTT TTA TTT GCA GGC TTT TCG TTA AGT CTA GTT GCA CAG      1945
216 Lys Asn Ala Leu Leu Phe Ala Gly Phe Ser Leu Ser Leu Val Ala Gln
217          500          505          510
219 TCA GTT AGT GCA CAA GAA GCA AAA CAG CCT GAA AAA GAA GAA AAA GAT      1993
220 Ser Val Ser Ala Gln Glu Ala Lys Gln Pro Glu Lys Glu Glu Lys Asp
221          515          520          525
223 GTT GAG GTG ATT TTG GTA TCG GCA CAA AAG CGT GAG CAA GCG CTT AAA      2041
224 Val Glu Val Ile Leu Val Ser Ala Gln Lys Arg Glu Gln Ala Leu Lys
225 530          535          540          545
227 GAA GTG CCT GTA TCA ATT GAA GTT ATT CAA GGC CTT CTA GA      2085
228 Glu Val Pro Val Ser Ile Glu Val Ile Gln Gly Asp Leu Leu
229          550          555
232 (2) INFORMATION FOR SEQ ID NO: 2:
234 (i) SEQUENCE CHARACTERISTICS:
235 (A) LENGTH: 559 amino acids
236 (B) TYPE: amino acid
237 (D) TOPOLOGY: linear
239 (ii) MOLECULE TYPE: protein
241 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
243 Met Arg Leu Tyr Phe Arg Lys Leu Trp Leu Thr Asn Leu Phe Leu Gly
244 1          5          10          15
245 Gly Ala Leu Ala Ser Ser Ala Ala Ile Gly Ala Val Ser Pro Lys Thr
246          20          25          30
247 Tyr Lys Asp Ala Asp Phe Tyr Val Ala Pro Thr Gln Gln Asp Val Asn
248          35          40          45
249 Tyr Asp Leu Val Asp Asp Phe Gly Ala Asn Gly Asn Asp Thr Ser Asp
250          50          55          60
251 Asp Ser Asn Ala Leu Gln Arg Ala Ile Asn Ala Ile Ser Arg Lys Pro
252 65          70          75          80
253 Asn Gly Gly Thr Leu Leu Ile Pro Asn Gly Thr Tyr His Phe Leu Gly
254          85          90          95
255 Ile Gln Met Lys Ser Asn Val His Ile Arg Val Glu Ser Asp Val Ile
256          100          105          110
257 Ile Lys Pro Thr Trp Asn Gly Asp Gly Lys Asn His Arg Leu Phe Glu
258          115          120          125
259 Val Gly Val Asn Asn Ile Val Arg Asn Phe Ser Phe Gln Gly Leu Gly
260          130          135          140

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/988,202

DATE: 01/15/2002

TIME: 19:00:49

Input Set : N:\Crf3\RULE60\09988202.raw

Output Set: N:\CRF3\01152002\I988202.raw

```

263 Asn Gly Phe Leu Val Asp Phe Lys Asp Ser Arg Asp Lys Asn Leu Ala
264 145 150 155 160
265 Val Phe Lys Leu Gly Asp Val Arg Asn Tyr Lys Ile Ser Asn Phe Thr
266 165 170 175
267 Ile Asp Asp Asn Lys Thr Ile Phe Ala Ser Ile Leu Val Asp Val Thr
268 180 185 190
269 Glu Arg Asn Gly Arg Leu His Trp Ser Arg Asn Gly Ile Ile Glu Arg
270 195 200 205
271 Ile Lys Gln Asn Asn Ala Leu Phe Gly Tyr Gly Leu Ile Gln Thr Tyr
272 210 215 220
273 Gly Ala Asp Asn Ile Leu Phe Arg Asn Leu His Ser Glu Gly Gly Ile
274 225 230 235 240
275 Ala Leu Arg Met Glu Thr Asp Asn Leu Leu Met Lys Asn Tyr Lys Gln
276 245 250 255
277 Gly Gly Ile Arg Asn Ile Phe Ala Asp Asn Ile Arg Cys Ser Lys Gly
278 260 265 270
279 Leu Ala Ala Val Met Phe Gly Pro His Phe Met Lys Asn Gly Asp Val
280 275 280 285
281 Gln Val Thr Asn Val Ser Ser Val Ser Cys Gly Ser Ala Val Arg Ser
282 290 295 300
283 Asp Ser Gly Phe Val Glu Leu Phe Ser Pro Thr Asp Glu Val His Thr
284 305 310 315 320
285 Arg Gln Ser Trp Lys Gln Ala Val Glu Ser Lys Leu Gly Arg Gly Cys
286 325 330 335
287 Ala Gln Thr Pro Tyr Ala Arg Gly Asn Gly Gly Thr Arg Trp Ala Ala
288 340 345 350
289 Arg Val Thr Gln Lys Asp Ala Cys Leu Asp Lys Ala Lys Leu Glu Tyr
290 355 360 365
291 Gly Ile Glu Pro Gly Ser Phe Gly Thr Val Lys Val Phe Asp Val Thr
292 370 375 380
293 Ala Arg Phe Gly Tyr Asn Ala Asp Leu Lys Gln Asp Gln Leu Asp Tyr
294 385 390 395 400
295 Phe Ser Thr Ser Asn Pro Met Cys Lys Arg Val Cys Leu Pro Thr Lys
296 405 410 415
297 Glu Gln Trp Ser Lys Gln Gly Gln Ile Tyr Ile Gly Pro Ser Leu Ala
298 420 425 430
299 Ala Val Ile Asp Thr Thr Pro Glu Thr Ser Lys Tyr Asp Tyr Asp Val
300 435 440 445
301 Lys Thr Phe Asn Val Lys Arg Ile Asn Phe Pro Val Asn Ser His Lys
302 450 455 460
303 Thr Ile Asp Thr Asn Thr Glu Ser Ser Arg Val Cys Asn Tyr Tyr Gly
304 465 470 475 480
305 Met Ser Glu Cys Ser Ser Ser Arg Trp Glu Arg Met Lys Gly Val Ser
306 485 490 495
307 Thr Lys Asn Ala Leu Leu Phe Ala Gly Phe Ser Leu Ser Leu Val Ala
308 500 505 510
310 Gln Ser Val Ser Ala Gln Glu Ala Lys Gln Pro Glu Lys Glu Glu Lys
311 515 520 525
312 Asp Val Glu Val Ile Leu Val Ser Ala Gln Lys Arg Glu Gln Ala Leu

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/988,202

DATE: 01/15/2002

TIME: 19:00:50

Input Set : N:\Crf3\RULE60\09988202.raw

Output Set: N:\CRF3\01152002\I988202.raw

L:35 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:36 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]